DOI:
10.1016/j.jaac.2019.01.021

Document Version
Peer reviewed version

Neuroimaging Evidence for Right Orbitofrontal Cortex Differences in Adolescents With Emotional and Behavioral Dysregulation

RH = Brain Correlates of Dysregulation

Philip A. Spechler, MA, Bader Chaarani, PhD, Catherine Orr, PhD, Scott Mackey, PhD, Stephen T. Higgins, PhD, Tobias Banaschewski, MD, PhD, Arun L.W. Bokde, PhD, Uli Bromberg, PhD, Christian Büchel, MD, Erin Burke Quinlan, PhD, Patricia J. Conrod, PhD, Sylvane Desrivières, PhD, Herta Flor, PhD, Vincent Frouin, PhD, Penny Gowland, PhD, Andreas Heinz, MD, PhD, Bernd Ittermann, PhD, Jean-Luc Martinot, MD, PhD, Frauke Nees, PhD, Dimitri Papadopoulos Orfanos, PhD, Luise Poustka, MD, Juliane H. Fröhner, MSc, Michael N. Smolka, MD, Henrik Walter, MD, PhD, Robert Whelan, PhD, Gunter Schumann, MD, Hugh Garavan, PhD, Robert R. Althoff, MD, PhD, the IMAGEN Consortium

Editorial
Supplemental Material

Accepted April 11, 2019

This article was reviewed under and accepted by Dr. Argyris Stringaris, MD, PhD.

Mr. Spechler and Drs. Chaarani, Higgins, Garavan, Althoff, Orr, and Mackey are with the University of Vermont, Burlington. Mr. Spechler and Drs. Chaarani, Higgins, Garavan, and Althoff are also with the Vermont Center on Behavior and Health, University of Vermont, Burlington. Drs. Banaschewski, Nees, and Flor are with the Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. Dr. Flor is also with the School of Social Sciences, University of Mannheim, Mannheim, Germany. Dr. Bokde is with the School of Medicine and Trinity College Institute of Neuroscience, Trinity College Dublin, Ireland. Drs. Bromberg and Büchel are with the University Medical Centre Hamburg-Eppendorf, Hamburg, Germany. Drs. Quinlan, Desrivières, and Schumann are with the Centre for Population Neuroscience and Stratified Medicine (PONS) and MRC-SGDP Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK. Dr. Conrod is with the Universite de Montreal, CHU Ste Justine Hospital, Canada. Drs. Frouin and Orfanos are with NeuroSpin, CEA, Université Paris-Saclay, France. Dr. Gowland is with the Sir Peter Mansfield Imaging Centre School of Physics and Astronomy, University of Nottingham, University Park, UK. Drs. Heinz and Walter are with Campus Charité Mitte, Charité, Universitätsmedizin Berlin, Germany. Dr. Ittermann is with Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany. Dr. Martinot is with the Institut National de la Santé et de la Recherche Médicale, INSERM Unit 1000 "Neuroimaging & Psychiatry", University Paris Sud – University Paris Saclay, France. Dr. Poustka is with the University Medical Centre Göttingen, Germany, and the Clinic for Child and Adolescent Psychiatry, Medical University of Vienna, Austria. Ms. Fröhner and Dr. Smolka are with Technische Universität Dresden, Germany. Dr. Whelan is with the School of Psychology and Global Brain Health Institute, Trinity College Dublin, Ireland.

This work received support from the following sources: the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain function and psychopathology) (LSHM-CT-2007-037286), the Horizon 2020 funded ERC Advanced Grant 'STRATIFY' (Brain network based stratification of reinforcement-related disorders) (695313), ERANID (Understanding the Interplay between Cultural, Biological and Subjective Factors in Drug Use Pathways) (PR-ST-0416-10004), BRIDGET (JPND: BRain Imaging, cognition Dementia and next generation GEnomics) (MR/N027558/1), the FP7 projects IMAGEMEND (602450; IMAging GEnetics for MENtal Disorders) and MATRICS (603016), the Innovative Medicine Initiative Project EU-AIMS (115300-2), the Medical Research Council Grant 'c-VEDA' (Consortium on Vulnerability to Externalizing Disorders and Addictions) (MR/N000390/1), the Swedish Research Council FORMAS, the Medical Research Council, the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, the Bundesministerium für Bildung und Forschung (BMBF grants 01GS08152; 01EV0711; eMED SysAlc01ZX1311A; Forschungsnetz AERIAL 01EE1406A, 01EE1406B), the Deutsche Forschungsgemeinschaft (DFG grants SM 80/7-2, SFB 940/2), and the Medical Research Foundation and Medical Research Council (grant MR/R00465X/1). Further support was provided by grants from: ANR (project AF12-NEUR0008-01-WM2NA and ANR-12-SAMA-0004), the Fondation de France, the Fondation pour la Recherche Médicale, the Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris Sud University IDEX 2012; the National Institutes of Health, Science Foundation Ireland (16/ERCD/3797), U.S.A. (Axon, Testosterone and Mental Health during Adolescence; RO1 MH085772-01A1), and by NIH Consortium grant U54 EB020403, supported by a cross-NIH alliance that funds Big Data to Knowledge Centres of Excellence. In addition, Drs. Garavan and Althoff are supported by P20GM103644 (PI: Stephen T. Higgins), Agency: NIGMS Vermont Center on Behavior and Health.

The authors wish to thank all members of the IMAGEN Consortium: Tobias Banaschewski MD, PhD, Heidelberg University; Gareth Barker, PhD, King's College London; Arun L.W. Bokde, PhD, Trinity College Dublin; Uli Bromberg, Dipl-Psych, University Medical Centre Hamburg-Eppendorf; Christian Büchel, MD, University Medical Centre Hamburg-Eppendorf; Erin Burke Quinlan, PhD, King's College London; Sylvane Desrivières, PhD, King's College London; Herta Flor, PhD, Heidelberg University and University of Mannheim; Vincent Frouin, PhD, Commissariat à l'Energie Atomique; Hugh Garavan, PhD, University of Vermont; Penny Gowland, PhD, University of Nottingham; Andreas Heinz, MD, PhD, Charité, Universitätsmedizin Berlin; Bernd Ittermann, PhD, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig and Berlin; Jean-Luc Martinot, MD, PhD, University Paris Sud, University Paris Descartes and Maison de Solenn; Marie-Laure Paillère Martinot, MD, PhD, University Paris Sud, University Paris Descartes and Maison de Solenn; Eric Artiges, MD, PhD, University Paris Sud, University Paris Descartes and Orsay Hospital; Herve Lemaitre, PhD, University Paris Sud and University Paris Descartes; Frauke Nees, PhD, Heidelberg University; Dimitri Papadopoulos Orfanos, PhD, Commissariat à l'Energie Atomique; Tomáš Paus, MD, PhD, University of Toronto; Luise Poustka, MD, Heidelberg University and Medical University of Vienna; Michael N. Smolka, MD, Technische Universität Dresden; Nora C. Vetter, PhD,

Technische Universität Dresden; Sarah Jurk, Dipl-Psych, Technische Universität Dresden; Eva Mennigen, MD, Technische Universität Dresden; Henrik Walter, MD, PhD, Charité, Universitätsmedizin Berlin; Robert Whelan, PhD, University College Dublin; Gunter Schumann, MD, King's College London. https://imagen-europe.com/

The authors thank Alexandra Ivanciu, BA, of the University of Vermont, for her assistance in preparing this manuscript.

Disclosure: Dr. Banaschewski has served as an advisor or consultant to Bristol-Myers Squibb, Desitin Arzneimittel, Eli Lilly and Co., Medice, Novartis, Pfizer, Shire, UCB, and Vifor Pharma. He has received conference attendance support, conference support, or speaking fees from Eli Lilly and Co., Janssen McNeil, Medice, Novartis, Shire, and UCB. He has been involved in clinical trials conducted by Eli Lilly and Co., Novartis, and Shire. The present work is unrelated to these relationships. Dr. Althoff is formerly employed, in part, by the nonprofit Research Center for Children, Youth, and Families. He has received grant or research support from the National Institute of Mental Health, the National Institute of General Medical Sciences, the National Institute on Drug Abuse, the Klingenstein Third Generation Foundation, and the Marcus Autism Center. He has served on the editorial board of Child Psychiatry and Human Development and as consulting editor of the Journal of Clinical Child and Adolescent Psychology. He has received honoraria from Oakstone Medical Publishing, Massachusetts General Hospital Psychiatry Academy, and Frontline Medical Communications, Inc. He is a partner of WISER Systems, LLC. Drs. Chaarani, Orr, Mackey, Higgins, Bokde, Bromberg, Büchel, Quinlan, Conrod, Desrivières, Flor, Frouin, Gowland, Heinz, Ittermann, Martinot, Nees, Orfanos, Poustka, Smolka, Walter, Whelan, Schumann, and Garavan, Mr. Spechler, and Ms. Fröhner report no biomedical financial interests or potential conflicts of interest.

Correspondence to Philip A. Spechler, MA, 1 So.Prospect St. Burlington, VT 05405; e-mail: philip.spechler@uvm.edu.

Abstract:

Objective: To characterize the structural and functional neurobiology of a large group of adolescents exhibiting a behaviorally and emotionally dysregulated phenotype.

Method: Age 14 adolescents from the IMAGEN study were investigated. Latent class analysis (LCA) on the Strengths and Difficulties Questionnaire (SDQ) was used to identify a class of individuals with elevated behavioral and emotional difficulties ("dysregulated"; n=233) who were compared to a matched sample from a low symptom class (controls, n=233). Whole-brain gray matter volume (GMV) images were compared using a general linear model with 10,000 random label permutations. Regional GMV findings were then probed for functional differences from three functional magnetic resonance imaging (fMRI) tasks. Significant brain features then informed mediation path models linking the likelihood of psychiatric disorders (*DSM-IV*) with dysregulation.

Results: Whole-brain differences were found in the right orbitofrontal cortex (R.OFC; $p<.05$; $k=48$), with dysregulated individuals exhibiting lower GMV. The dysregulated group also exhibited higher activity in this region during successful inhibitory control ($F_{1,429}=7.53$, $p<.05$). Path analyses indicated significant direct effects between the likelihood of psychopathologies and dysregulation. Modeling the R.OFC as a mediator returned modest partial effects, suggesting the path linking the likelihood of an anxiety or conduct disorder diagnoses to dysregulation is partially explained by this anatomical feature.

Conclusion: A large sample of dysregulated adolescents exhibited lower GMV in the R.OFC relative to controls. Dysregulated individuals also exhibited higher regional activations when exercising inhibitory control at performance levels comparable to controls. These findings

suggest a neurobiological marker of dysregulation and highlight the role of the R.OFC in impaired emotional and behavioral control.

Key words: dysregulation, SDQ, orbitofrontal cortex, adolescence, VBM

Introduction:

Adolescents exhibiting severe difficulties regulating behavior and emotion are commonly referred for psychiatric evaluation but are difficult to classify into discrete diagnostic categories, with "comorbidity" being the rule rather than the exception in child psychiatry. Previous labels for these dysregulated children included severe mood dysregulation (SMD) or irritability[1] with the acknowledgement that these individuals will likely meet diagnostic criteria for other disorders. Recently, "disruptive mood dysregulation disorder" (DMDD)[2] was added to the *DSM-5* to better classify dysregulated children. Research indicates the prevalence of dysregulation is between 0.8 and 3.3%, with particularly high co-occurrence with externalizing and internalizing disorders.[3,4] As individuals with a singular diagnosis may be thought of as behaviorally *or* emotionally dysregulated, it is specifically the individuals with a set of difficulties spanning both behavioral and emotional domains who need to be studied further. Considering the addition of this disorder into the *DSM*, and research showing the functional outcomes of dysregulated youths are strikingly poor,[5] it is imperative to identify the neurobiological correlates of dysregulation. Characterizing the pathophysiological substrates will help inform dysregulation nosology, provide diagnostic biomarkers, and help inform targeted treatment methods.

The NIMH recently advocated the Research Domain Criteria (RDoc), which hypothesizes psychiatric problems coexist on a spectrum of severity with symptoms that cut across discrete diagnostic categories. Therefore, in this report using a large dataset of adolescents (the IMAGEN study),[6] we adopted a latent class analysis (LCA) approach to the Strengths and Difficulties Questionnaire (SDQ)[7] to identify groups of individuals endorsing similar patterns of behavioral and emotional problems. The result of an SDQ-LCA provides class groupings, as well as dimensional characteristics of emotional and behavioral problems hypothesized to contain

varying patterns of symptomatology that resist discrete categorization. One class is specifically hypothesized to comprise a profile analogous to DMDD. In other words, in the absence of DMDD diagnoses, we hypothesized a class of individuals exhibiting a profile in line with a dysregulation phenotype. Although measurement of a dysregulation profile is a major challenge in the field,[8–10] the intent of our investigation is to characterize the neural correlates of dysregulation as defined by one measurement method (among a suite of others).[11,12]

Structural neuroimaging, and specifically, voxel-based morphometry (VBM), has been used to study many psychiatric constructs across stages of development. VBM allows the researcher to measure the volumes of the major tissue types of the brain,[13] thus providing a neurobiological framework to closely study a behavioral profile of interest. VBM has informed many adolescent psychiatric disorders related to dysregulation including anxiety,[14] depression,[15] and conduct disorder.[16] Regarding previous structural neuroimaging studies of dysregulation, Adleman and colleagues used VBM to uncover differences among children with SMD, bipolar disorder (BP), and controls, with the SMD group exhibiting the lowest gray matter volume (GMV) in bilateral pre-supplemental motor area, right insula and dorsolateral prefrontal cortex.[17] Additionally, Gold and colleagues used VBM to study youths with anxiety, BP, ADHD, and DMDD, compared to controls. Gold found GMV differences specific to, and across psychiatric disorders, with dysregulated participants exhibiting lower GMV in the right dorsolateral and superior frontal cortex.[18] Therefore, for our primary analysis using whole-brain VBM data, we hypothesized dysregulated individuals would exhibit lower GMV relative to controls in cortical regions implicated in regulatory processes such as the bilateral insula, right-sided dorsolateral prefrontal cortex, and ventromedial/orbitofrontal cortex. [17,19]

Regions uncovered from the primary anatomical analysis can be used as regions of interest in post-hoc analyses on the fMRI data from the IMAGEN study. These post-hoc analyses broadly test the hypothesis that differences in brain structure yields differences in brain function. Interrogating both structure and function with the same dataset maximizes the information gained about the neurobiological characteristics of dysregulation, and captures the brain's trait-like features measured via structural neuroimaging, and state-like features measured during functional task demands. Follow-up analyses on neuroimaging data can also be used to explain the relationship between candidate comorbidity diagnoses[3] and dysregulation. For instance, an identified neurobiological correlate of dysregulation can be modeled as a mediator in a path analysis linking the likelihood of a psychiatric disorder to dysregulation. In doing so, we test the hypothesis that the brain mediates the relationship between a disorder and dysregulation in some linear fashion. As we only probe data from the age 14 assessment of the IMAGEN study, these mediation models infer correlation and not causation.

Method:

Participants were drawn from the IMAGEN study of adolescent development.[6] Comprehensive study details are available in the online Standard Operating Procedures (https://imagen-europe.com/). The IMAGEN study conformed to the ethical standards outlined by the Declaration of Helsinki and was approved by ethics committees at each site including King's College, London; Central Institute of Mental Health, Mannheim; Charite, Universitatsmedizin Berlin; University Medical Center Hamburg-Eppendorf; University of Nottingham; Trinity College Dublin; Institut National de la Sante et de la Recherche Medicale, Orsay. After description of the study to participants and their parents, written informed consent was obtained. Individuals who provided assent were assessed at age 14. For this report, all data were taken from the baseline assessment only (age 14). Participants with SDQ data (N=2,126) were used as the starting sample of the analysis (Age *M*=14.56, *SD*=.44; Female individuals =1,081, 51%). Selected participants from the LCA analysis were then drawn from the sample who received an anatomical scan with GMV images passing quality control (N=2,024).

Strengths and Difficulties Questionnaire

The SDQ is 25-item instrument designed to characterize children across five domains including emotional symptoms, conduct problems, hyperactive behavior, peer problems, and prosocial behaviors.[7] Hence, the SDQ is especially suited to capture both the behavioral and emotional features related to dysregulation. Furthermore, the SDQ is widely used and has been shown to predict psychiatric diagnoses later in life.[20] Data included in the analysis were from the parent reporting on their child's behavior in the past six months. SDQ data from N=2,126 participants were used in the latent class analysis.

Each SDQ item is measured on an ordinal scale: 0=Not True, 1=Somewhat True, 2=Certainly True. While the majority of the instrument is negatively valenced (e.g., "Often unhappy", "Often lies or cheats"), the few positively valenced items are reversed coded with the exception of the entire prosocial domain. Therefore, higher values within the emotional, conduct, hyperactive or peer domain reflect difficulties, whereas higher values within the prosocial domain reflect strengths. For the input to the latent class analysis, positively valenced items from the prosocial domain were recoded to match the overall pattern of the instrument.

Previous investigators have reported using the SDQ-Dysregulation Profile (SDQ-DP) to measure the dysregulation construct based on the sum of five proposed items.[21] Rather than imposing the recommended SDQ-DP cutoff of scores ≥5 as dysregulated, we used a data driven approach to characterize individuals based on patterns of similar problem behaviors. And while the SDQ-DP is based on five SDQ items spanning behavioral and affective problems, youths who score high on only the behavioral items may be categorized as dysregulated despite scoring low on the emotional items. The use of latent class analysis is hypothesized to overcome this limitation by identifying a class of individuals who are most likely to exhibit co-occurring behavioral and affective problems. Nonetheless, the SDQ-DP was calculated and compared to the class probabilities returned from the latent class analysis.

Latent Class Analysis

Latent class analysis (LCA) is an example of a mixture model used to estimate group membership of latent constructs. LCA is robust to the categorical data format of the SDQ and assigns probability scores to each participant reflecting the likelihood of class membership. Participants were categorized into the class with the highest probability score.

Latent class models were estimated using the software Mplus via an EM algorithm. Model comparison was performed on analyses returning 1-class through 7-class solutions. The best-fitted model was identified using multiple measures of fit. The Bayesian Information Criterion (BIC) is a goodness-of-fit index that penalizes models with more classes. Lower BIC values indicate more parsimonious models. Because standard loglikelihood tests are biased in this analytic environment, two other examinations were used to compare a K class model to a K-1 class model, the Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMRT) and the bootstrap likelihood ratio test (BLRT). In each case, significance comparing a K class model to a K-1 class model indicates additional information is provided. If it is not significant, then the K-1 class model can be accepted. In addition, models with higher entropy (closer to 1) indicate a clearer delineation of classes.[22] In this analysis, all indices other than the BLRT (which was not discriminating) indicated a 5-class model fit (see Table 1). These classes were then used to identify two groups of interest, a dysregulated group, and a low symptom comparison group (controls). Group identification was determined based on their respective patterns of item endorsement as further explained below.

Structural Neuroimaging Methods

Across the eight acquisition sites, participants were scanned on 3T MRI scanners from various manufacturers (Phillips, General Electric, Bruker, and Siemens). Standardization and quality assurance efforts were made to insure all sites used the same MRI acquisition parameters and yielded comparable data. High-resolution anatomical magnetic resonance images were acquired, including a 3D T1-weighted magnetization prepared gradient echo sequence based on the ADNI protocol (http://adni.loni.usc.edu). The structural image was collected for nine minutes

using the following parameters: TR=2300ms; TE=2.8ms; flip angle=8°; matrix size=240x256; voxel resolution=1.1mm^3; and 160 contiguous slices at a thickness of 1.1mm.

Whole-brain gray matter volume (GMV) images were generated using optimized voxel-based morphometry procedures in SPM8.[13] High-resolution anatomical magnetic resonance images were acquired, including a 3D T1-weighted magnetization prepared gradient echo sequence based on the ADNI protocol (http://adni.loni.usc.edu). Structural MRI preprocessing included segmenting and normalizing the images into Montreal Neurological Institute template space. The gray matter segmentation images were then modulated to obtain volumetric images, rather than tissue concentration images. N=2,024 participants received a structural MRI and had GMV images passing quality control.

In preparation for the between-group GMV comparison, variables potentially influencing adolescent GMV (age, sex, site of imaging acquisition, handedness, puberty status,[23] verbal and performance IQ,[24] and total GMV) were partialled out of the images. To do so, all participants from the baseline IMAGEN sample with preprocessed GMV images (N=2,024) were submitted to a multiple regression with only the confounding variables included in the design matrix. The residual GMV image for each participant was then used in the permutation test described below. This procedure was used because including nuisance covariates in the permutation analysis prohibitively increased computation time.

In light of recent criticisms related to the proper correction for multiple comparisons in neuroimaging research,[25] permutation analyses have been advocated as a non-parametric approach to closely control the number of false-positives in a statistical analysis.[26] Here, we used a random label permutation test applied to the residual output of the aforementioned nuisance regression. Each participant's group membership was randomly shuffled and a whole-brain two-

group *t*-test using a general linear model was fitted to the residualized images. Random label shuffling was repeated 10,000 times, thus building a null distribution at each voxel, to which the originally labeled results were compared. Threshold-free cluster enhancement correction (TFCE)[27] was then used to control the family-wise error rate for identifying clusters of residual gray matter that exhibit significant group differences. Regions of interest (ROI) surviving a TFCE corrected $\alpha < .05$ were then probed for fMRI group differences, as well as being modeled as the mediator in candidate path analyses linking the likelihood of psychopathology to dysregulation. Permutation analyses were conducted using FSL's Permutation Analysis of Linear Models[26] on the University of Vermont's Advanced Computing Core.

Functional Neuroimaging Methods

Three fMRI tasks commonly used in psychiatric neuroimaging were administered, including the stop signal task, monetary-incentive delay task, and a face-processing task. The stop signal (inhibitory control) task requires participants to inhibit a prepotent motor response.[28] Motor inhibitory control performance during this task is commonly measured using the stop signal reaction time (SSRT), an estimate of the speed of the inhibitory process, calculated from the average latency period between the "go" and "stop signal" during successful inhibition trials.[29] The monetary-incentive delay task measures the processing of both anticipation and receipt of monetary rewards.[30] The face-processing task involves the passive viewing of angry faces, neutral faces, and control images.[31] See Supplement 1, available online, for full details on the fMRI tasks.

All fMRI data were submitted to standard preprocessing methods and whole-brain contrast images specific to each task were estimated using a general linear model (see Supplement 1,

available online, for fMRI processing details). Specifically, unsuccessful and successful inhibitory control, monetary reward anticipation and receipt, angry faces, neutral faces, and the differential activation for angry minus neutral faces, were each used to explore any functional differences between the groups. For each contrast image, the mean value within a region of interest (ROI) was extracted and analyzed using two-group ANCOVA models with a Bonferroni corrected alpha based on the number of contrasts tested for each task modality.

Likelihood of Psychiatric Diagnoses

Psychopathology was determined using the Developmental and Well-Being Assessment (DAWBA; http://www.dawba.info/a0.html), a set of computer-administered interviews, questionnaires, and ratings generating *DSM-IV* psychiatric diagnoses for ages 5-17. Based on the child and parent responses, a computer algorithm generates scores to predict the likelihood of meeting criteria for *DSM-IV* diagnoses ("band scores"). These band scores range from 1 to 5, representing a probability of <0.1% to >70%. DAWBA band scores have been shown to yield prevalence estimates that broadly compare to clinician-rated diagnoses.[20]

Mediation Analyses

Mediation was conducted in Mplus using a robust weighted least squares estimator to estimate direct and indirect effects, with bias-corrected 95% confidence intervals generated from 1000 bootstrapped samples. The use of bootstrapping the indirect effects is a more powerful method of inferring mediation compared to the traditional five-step procedure.[32] The independent variables for the five separate mediation analyses included the full range of DAWBA band scores, representing the likelihood of receiving a *DSM-IV* diagnosis for anxiety, depression,

conduct disorder (CD), oppositional defiance disorder (ODD), and attention deficit hyperactivity disorder (ADHD). These five constructs were informed by Copeland and colleagues who assessed the prevalence rates of DMDD comorbidity with these disorders.[3] Furthermore, these disorders broadly capture the co-occurring internalizing and externalizing problems exhibited by dysregulated individuals.

The identified GMV features were modeled as a mediator between each band score and the binary dysregulation status as the dependent variable. Hence, models were constructed to test the hypothesis that the underlying neurobiology influences the relationship between a related psychiatric disorder and the dysregulated phenotype. As the initial neuroimaging analysis here tests for a biomarker of dysregulation in isolation, follow up path analyses assessed the involvement of brain structure with dysregulation in the context of affiliated psychiatric diagnoses reported by Copeland and colleagues. Any significant indirect effects uncovered by these path models provides evidence indicating the correlation between the likelihood of a related disorder and being dysregulated is driven, in part, through changes in focal brain structure.

Results:

Latent Class Analysis Results

The best-fitting LCA model returned a five-class solution (see Table 1 for fit statistics). Here, we describe each class and offer a label to characterize their profile. Class 1, the "defiant class" (18% of the sample), contained individuals with low prosocial traits and slightly elevated conduct problems and hyperactivity. Class 2, the "emotional difficulties" class (16% of the sample), contained individuals with the highest emotional difficulties. Class 3, the "dysregulated class" (12% of the sample), contained individuals with very high levels of difficulties across all five domains. Class 4, the "hyperactive class" (25% of the sample), contained individuals with elevated hyperactivity. Class 5, the "low symptom class" (29% of the sample), contained individuals with very low levels of problem behaviors across all domains. And while class 5 is labeled "low symptoms", we note that the defiant, emotional difficulties, and hyperactive classes also exhibit low symptoms on domains outside of their problem areas. These findings are consistent with studies reporting high prevalence rates of any level of psychiatric symptomatology in adolescence.[33] See Figure 1 for the average item-endorsement for each class and Table 2 for the five SDQ summary scores for each class.

While other classes exhibited elevations in a single domain (i.e., emotional difficulties class; hyperactive class), the dysregulated class distinctly exhibited co-occurring behavioral and affective problems. These individuals exhibited the highest probability of endorsing conduct problems, hyperactivity, peer problems, and the second highest probability of endorsing emotional problems (closely following the emotional difficulties class), and prosocial problems (closely following the defiant class). The low symptom class (the largest sample) was selected as the comparison group as they exhibited the lowest probability of endorsing all problematic

behaviors. See Table S1, available online, for comparison of the dysregulated class to the full sample on descriptive characteristics.

Next, the SDQ-DP was calculated to compare to the LCA results using bivariate correlations between the SDQ-DP sum score and the probability of each class membership. Results indicated the SDQ-DP was most positively associated with the dysregulated class (r_{2126} = .61, p<.001) and most negatively associated with the low symptom class (r_{2126} = -.44, p <.001), thus providing support for the dysregulated phenotype captured by class 3 and the low symptom group captured by class 5. However, these correlations may be inflated as both measures were estimated from similar items on the same dataset. Nonetheless, Holtmann and colleagues report correlations between their SDQ-DP and Child Behavior Checklist-Dysregulation Profile (CBCL-DP) binary score at r=.45 and CBCL-DP sum score at r=.75. Therefore, the LCA results reported here are in line with these other measurement instruments.

There were 184 participants included in the LCA who did not provide anatomical scan data (for reasons including failing quality control, MRI safety issues, etc.). However, chi-square tests indicated there was no difference in LCA class membership in this subsample relative to the larger sample with anatomical scan data ($\mathbf{X}^2_{4,2126}$ = 2.2 , p > .05). Thus, we do not believe there was any skew in the LCA class assignment by the participants who did not provide anatomical scan data.

Neuroimaging Sample Characteristics

Of the 261 dysregulated and 613 low symptom comparison individuals identified from the LCA, 233 dysregulated and 564 comparison individuals provided useable GMV data. For the sample of 233 dysregulated individuals, an equal size subset of comparison individuals was

selected from the low symptom class. This control group was pseudo-randomly selected so as to match to the dysregulated group by containing an equal number of male and female individuals who showed no differences on total GMV, pubertal development, performance and verbal IQ, or age, and contained similar distributions for handedness and site of acquisition (see Table S2, available online, for group comparisons). And while site was included in the initial nuisance regression of the full IMAGEN dataset, it is difficult to precisely account for site when there are unequal representations at each site. Hence, a pseudo-random sampling of the two groups was performed to identify a subsample of individuals with equal representations at each site. Results using this perfectly site-matched subsample were consistent with the main findings reported below. See Supplement 1, available online, for more information.

Whole-brain Residual Gray Matter Volume Comparison

After running a two-sample *t*-test using a general linear model with 10,000 random label permutations, a single cluster survived TFCE-correction for multiple comparisons ($P_{FWE\text{-}corr}<.05$, $k=48$ voxels). This cluster was found in the right orbitofrontal cortex (OFC), center of mass at (MNI: 24, 30, -16), spanning the orbital sulcus with extent into the posterior orbital gyrus. In this region, dysregulated individuals exhibited lower residual GMV relative to their peers with low symptoms (see Figure 2).

Laterality Test

As only one hemisphere survived strict correction, and there is growing interest in prefrontal asymmetry, a contralateral region of interest analysis was performed post-hoc. To perform this test, we translated the right-sided region of interest onto the left hemisphere and

extracted regional GMV for all subjects. Two-sample *t*-tests indicated the left OFC ROI yielded significant differences (L.OFC: t_{462} = -3.32, $p < 1.0\text{x}10^{-3}$), similar to the findings in the right OFC, albeit at a relatively lower magnitude of effect (R.OFC: t_{462} = -4.40, $p < 1\text{x}10^{-4}$).

fMRI Comparisons

The identified sample for GMV analyses (n=466) was selected on the basis of the quality of their anatomical image, meaning some of these participants did not have fMRI data available. See Table S3 and Supplement 1, available online, for full details regarding these reduced samples, and reasons for missingness. In preparation for the ROI-level between-group comparisons using the fMRI data, we first examined the amount of head motion in the images. For each subject, the mean framewise displacement (mean FD) was calculated for each of the three fMRI tasks. Based on prior developmental neuroimaging studies,[34] a head motion exclusionary criterion of mean FD > .9mm was used. For the stop signal task, 3 dysregulated participants were excluded. For the faces task, 1 dysregulated participant was excluded. For the MID task, 5 dysregulated and 1 low symptom participants were excluded. Importantly, these reduced samples for fMRI comparisons retained critical between-group similarities as the starting samples for anatomical comparisons. Chi-square (for categorical measures) and *t*-tests indicated that after excluding subjects, the reduced samples retained their best-matched characteristics and did not differ on age, sex, handedness, IQs, or total GMV ($p > .05$).

Data were then submitted to standard two-sample *t*-tests to determine any group differences in head motion for a given task. Results indicated that while mean FD did not exceed thresholds previously reported as problematic,[35,36] the dysregulated sample exhibited significantly more head motion during each fMRI acquisition (see Table S4, available online).

Although participants' head motion parameters were included in the design matrix during their fMRI contrast estimation, we also included mean FD as a covariate in the ROI-level between-group ANCOVA models.

For the stop signal task, results indicated a significant between-group difference during successful inhibitory control trials $F_{1,377}$=5.61, p_{corr}<.05, such that the dysregulated group showed higher activation (n=186, *M*=.15, *SD*=1.2) than the low symptom group (n=194, *M*=-.09, *SD*=.92). To ensure these findings were not driven by the difference in head motion, similar ANCOVA models were estimated on 5,000 pseudo-random subsamples of the data matched on head motion. Results were consistent, leading to a mean $F_{1,307}$=4.9, p<.05, suggesting the between-group difference on successful inhibitory control activations were not driven by head motion. See Supplement 1, available online, for more information.

Due to problems with the behavioral task performance adaptive algorithm, stop signal reaction time (SSRT) scores were available on only a subset of participants. A between-group comparison on those individuals with useable SSRT behavioral data (Dysregulated n=97; Controls n=107) yielded no significant differences on SSRT (t_{202}=0.38, p = .71). Given the reduced sample sizes of participants with SSRT data, no imputations were performed for SSRT, and it is unknown the degree to which the effects might generalize to the starting samples. No between-group activation differences were detected for unsuccessful inhibitory control trials, or on any of the remaining fMRI contrasts (reward and face processing tasks).

Mediation Analyses

The likelihood of having any of the five psychopathologies exhibited a significant total and direct effect with dysregulation, substantiating their relationship with dysregulation.[3] Bias-

corrected confidence intervals around the indirect effect of the right OFC GMV ROI indicated this region partially mediated the likelihood of an anxiety disorder diagnosis (c=.023, 95% CI [0.003, 0.043]) or conduct disorder diagnosis (c=.018, 95% CI [.003, .033]) to dysregulation status (see Figure 3). No significant indirect effects were detected to link the brain between the likelihood of depression, ODD, or ADHD with dysregulation. Additionally, regional fMRI brain activation during successful inhibitory control did not yield any significant mediation effects. See Table 3 for mediation model results.

Discussion:

We report that emotionally and behaviorally dysregulated adolescents exhibited lower GMV in the right OFC relative to their non-dysregulated peers. These findings were identified by a conservative permutation analysis between two large samples of closely matched groups. Secondary analyses indicated that within this same region, the dysregulated group exhibited higher functional brain activation when executing successful inhibitory control behaviors. These fMRI results provide some specificity on the psychological correlates of the GMV effect, such that the anatomical difference associated with dysregulation was related to inhibitory control but not to face or reward processing. Taken together, these results suggest dysregulation is characterized by differences in cortical regions involved with executive functioning. Lastly, the volume of the right OFC region partially mediated relationships between the likelihoods of an anxiety disorder and a conduct disorder diagnosis and dysregulation.

It is interesting that the right OFC was uncovered from a conservative whole-brain analysis and also exhibited differences on the stop signal task, as there is a body of research implicating the OFC in behavioral and emotional regulation. For example, previous research on the IMAGEN sample identified this region as participating in a network of brain activity during successful inhibitory control trials.[37] As the dysregulated and low symptoms groups exhibited similar task performance, the greater activity in the right OFC of the dysregulated group may reflect greater effort or cognitive resources needed to execute inhibitory behaviors equal to that of their peers. Therefore, dysregulation may be partly dependent on a neurobiological inhibitory control network compromised in its ability to efficiently regulate behavior.

The OFC is also putatively involved in integrating attention and emotion by assigning a signal of affective value to stimuli. Previous work using event-related potentials (ERP) during an

affective go/no-go task was conducted on children with co-occuring internalizing and externalizing disorders. One set of results identified higher ventral prefrontal activations during inhibitory control trials in children with poor self-regulatory abilities as measured via parent-child observations.[38] In a related treatment study of similar children, treatment success was characterized by attenuation of activation levels in the ventral prefrontal region during inhibitory control trials.[39] Hence, our findings are in line with these reports and suggest the OFC as both a potential biomarker and candidate region for targeted clinical interventions to help improve outcomes in children with dysregulated behavioral profiles.

In terms of the mediation results, the investigated psychopathologies all exhibited a significant and large total effect on dysregulation, indicating that the likelihood of having an internalizing or externalizing disorder was associated with an increased likelihood of being dysregulated. These findings are consistent with previous reports identifying similar patterns of comorbidity from three datasets of child psychopathology.[3] Moreover, the direct effects were also large, accounting for nearly 98% of the total effect for all disorders (see Table 3). Given these relationships, the significant partial mediation results are notable as little variance is left to be explained by the indirect paths. Yet despite these relatively weak indirect effects, the significant findings highlight the transdiagnostic nature of the right OFC region insofar as it is a mediator to dysregulation for the likelihood of anxiety and conduct disorder. Although a significant mediation was not observed for depression, ODD, or ADHD it would be incautious to conclude that the mediation effect has specificity for anxiety and conduct disorder as effects in similar directions were observed for depression (p=.065) and ODD (p=.070; see Table 3). On the whole, the data suggest a small but potentially important role for the OFC in linking internalizing and externalizing disorder to dysregulation. Lastly, we reiterate the path models should not be

misinterpreted as implying the likelihood of an anxiety, conduct disorder, or the brain feature caused dysregulation as the models are restricted to age 14 data only.

Limitations of this study include the lack of DSM-5 diagnostic measures, as it is unclear if the individuals contained in the dysregulation group meet DMDD diagnostic criteria. Future studies are needed to evaluate the degree to which the SDQ captures individuals who receive a *DSM-5* DMDD diagnosis following a clinical interview. Likewise, measurement studies are also needed to determine the correlation between popular measurement methods like SDQ-LCA, SDQ-DP and CBCL-DP, and their correlation with clinical ratings. Additionally, recent work taking a factor analytic approach to the SDQ has identified a dysregulation factor using just three of the five domains, omitting the prosocial and peer problem domains.[10] However, given that elevations in the CBCL Social Problems domain frequently accompany the CBCL-DP,[8] this approach risks omitting relevant features of the dysregulation construct.

In considering dysregulation measurement inconsistencies, differences in the precise brain region uncovered here with the previous regions uncovered by Adleman and by Gold and colleagues are likely attributed to differing measurement approaches. Nonetheless, the right-sided prefrontal anatomical finding is broadly consistent with these prior results. Although earlier fMRI studies of inhibitory control in dysregulation failed to detect significant group differences[41] this is likely due to our fMRI analysis, by design, being restricted to a single anatomically defined region of interest. Another important consideration in interpreting the present fMRI results and integrating them with past findings is the potential confounding role of head motion. Despite including mean framewise displacement as a covariate, ANCOVA models are generally unable to completely control for a significant between-group difference in that covariate. Confidence in our results comes from the 5,000 pseudo-random subsampling

procedure in which the group differences were recapitulated with subsamples chosen not to differ on head motion.

A caveat regarding the mediation results is that the path models were estimated using DAWBA band scores. As the DAWBA contains many skip rules leading some participants to "screen in" for extra items, these skips rules are sometimes related to high SDQ domain scores. Estimating paths between band scores and a binary dysregulation score determined via an SDQ-LCA consequently may contain a degree of circularity. Another limitation of the present study is the use of single informant data, although previous studies suggest agreements among multi-informants are generally low.[42] Finally, given the neurodevelopmental changes underway at age 14, it is unknown if the neuroanatomical difference identified here persists throughout the lifespan. Future longitudinal studies on dysregulated individuals are needed to determine the psychosocial and neurobiological antecedents of dysregulation, as well as the developmental effect of neural maturation on the persistence of dysregulation into late adolescence and adulthood.

References

1. Leibenluft E. Severe Mood Dysregulation, Irritability, and the Diagnostic Boundaries of Bipolar Disorder in Youths. *Am J Psychiatry*. 2011;168(2):129-142.

2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub; 2013.

3. Copeland WE, Angold A, Costello EJ, Egger H. Prevalence, Comorbidity, and Correlates of DSM-5 Proposed Disruptive Mood Dysregulation Disorder. *Am J Psychiatry*. 2013;170(2):173-179.

4. Dougherty LR, Smith VC, Bufferd SJ, et al. DSM-5 disruptive mood dysregulation disorder: correlates and predictors in young children. *Psychol Med*. 2014;44(11):2339-2350.

5. Copeland WE, Shanahan L, Egger H, Angold A, Costello EJ. Adult Diagnostic and Functional Outcomes of DSM-5 Disruptive Mood Dysregulation Disorder. *Am J Psychiatry*. 2014;171(6):668-674.

6. Schumann G, Loth E, Banaschewski T, et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol Psychiatry*. 2010;15(12):1128-1139.

7. Goodman R. The Strengths and Difficulties Questionnaire: A Research Note. *J Child Psychol Psychiatry*. 1997;38(5):581-586.

8. Althoff RR, Ayer LA, Rettew DC, Hudziak JJ. Assessment of dysregulated children using the Child Behavior Checklist: A receiver operating characteristic curve analysis. *Psychol Assess*. 2010;22(3):609-617.

9. Deutz MHF, Geeraerts SB, van Baar AL, Deković M, Prinzie P. The Dysregulation Profile in middle childhood and adolescence across reporters: factor structure, measurement invariance, and links with self-harm and suicidal ideation. *Eur Child Adolesc Psychiatry*. 2016;25(4):431-442.

10. Deutz MHF, Shi Q, Vossen HGM, et al. Evaluation of the Strengths and Difficulties Questionnaire-Dysregulation Profile (SDQ-DP). *Psychol Assess*. 2018;30(9):1174-1185.

11. Althoff RR, Verhulst F, Rettew DC, Hudziak JJ, van der Ende J. Adult Outcomes of Childhood Dysregulation: A 14-year Follow-up Study. *J Am Acad Child Adolesc Psychiatry*. 2010;49(11):1105-1116

12. Jordan P, Rescorla LA, Althoff RR, Achenbach TM. International Comparisons of the Youth Self-Report Dysregulation Profile: Latent Class Analyses in 34 Societies. *J Am Acad Child Adolesc Psychiatry*. 2016;55(12):1046-1053.

13. Ashburner J, Friston KJ. Voxel-Based Morphometry—The Methods. *NeuroImage*. 2000;11(6):805-821.

14. Radua J, van den Heuvel OA, Surguladze S, Mataix-Cols D. Meta-analytical Comparison of Voxel-Based Morphometry Studies in Obsessive-Compulsive Disorder vs Other Anxiety Disorders. *Arch Gen Psychiatry*. 2010;67(7).

15. Bora E, Fornito A, Pantelis C, Yücel M. Gray matter abnormalities in Major Depressive Disorder: A meta-analysis of voxel based morphometry studies. *J Affect Disord*. 2012;138(1-2):9-18.

16. Fairchild G, Passamonti L, Hurford G, et al. Brain Structure Abnormalities in Early-Onset and Adolescent-Onset Conduct Disorder. *Am J Psychiatry*. 2011;168(6):624-633.

17. Adleman NE, Fromm SJ, Razdan V, et al. Cross-sectional and longitudinal abnormalities in brain structure in children with severe mood dysregulation or bipolar disorder. *J Child Psychol Psychiatry*. 2012;53(11):1149-1156.

18. Gold AL, Brotman MA, Adleman NE, et al. Comparing Brain Morphometry Across Multiple Childhood Psychiatric Disorders. *J Am Acad Child Adolesc Psychiatry*. 2016;55(12):1027-1037.

19. Rogers JC, De Brito SA. Cortical and Subcortical Gray Matter Volume in Youths With Conduct Problems: A Meta-analysis. *JAMA Psychiatry*. 2016;73(1):64.

20. Goodman A, Goodman R. Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain. *J Child Psychol Psychiatry*. 2011;52(1):100-108.

21. Holtmann M, Becker A, Banaschewski T, Rothenberger A, Roessner V. Psychometric Validity of the Strengths and Difficulties Questionnaire-Dysregulation Profile. *Psychopathology*. 2011;44(1):53-59.

22. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *J Classif*. 1996;13(2):195-212.

23. Carskadon MA, Acebo C. A self-administered rating scale for pubertal development. *J Adolesc Health*. 1993;14(3):190-195.

24. Wechsler D. Wechsler intelligence scale for children–Fourth Edition (WISC-IV). *San Antonio TX Psychol Corp*. 2003.

25. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci*. 2016;113(28):7900-7905.

26. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *NeuroImage*. 2014;92:381-397.

27. Smith S, Nichols T. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. 2009;44(1):83-98.

28. Rubia K, Smith AB, Taylor E, Brammer M. Linear age-correlated functional development of right inferior fronto-striato-cerebellar networks during response inhibition and anterior cingulate during error-related processes. *Hum Brain Mapp*. 2007;28(11):1163-1177.

29. Logan GD, Cowan WB. On the ability to inhibit thought and action: A theory of an act of control. *Psychol Rev*. 1984;91(3):295-327.

30. Knutson B, Westdorp A, Kaiser E, Hommer D. FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*. 2000;12(1):20-27.

31. Grosbras M-H. Brain Networks Involved in Viewing Angry Hands or Faces. *Cereb Cortex*. 2005;16(8):1087-1096.

32. Hayes AF. Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Commun Monogr*. 2009;76(4):408-420.

33. Copeland W, Shanahan L, Costello EJ, Angold A. Cumulative prevalence of psychiatric disorders by young adulthood: a prospective cohort analysis from the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry*. 2011;50(3):252-261.

34. Silvers JA, Lumian DS, Gabard-Durnam L, et al. Previous Institutionalization Is Followed by Broader Amygdala–Hippocampal–PFC Network Connectivity during Aversive Learning in Human Development. *J Neurosci*. 2016;36(24):6420-6430.

35. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 2012;59(3):2142-2154.

36. Siegel JS, Power JD, Dubis JW, et al. Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum Brain Mapp*. 2014;35(5):1981-1996.

37. Whelan R, Conrod PJ, Poline J-B, et al. Adolescent impulsivity phenotypes characterized by distinct brain networks. *Nat Neurosci*. 2012;15(6):920-925.

38. Granic I, Meusel L-A, Lamm C, Woltering S, Lewis MD. Emotion regulation in children with behavior problems: Linking behavioral and brain processes. *Dev Psychopathol*. 2012;24(3):1019-1029.

39. Lewis MD, Granic I, Lamm C, et al. Changes in the neural bases of emotion regulation associated with clinical improvement in children with behavior problems. *Dev Psychopathol*. 2008;20(3).

40. Blair RJR. Psychopathy, frustration, and reactive aggression: The role of ventromedial prefrontal cortex. *Br J Psychol*. 2010;101(3):383-399.

41. Deveney CM, Connolly ME, Jenkins SE, et al. Neural recruitment during failed motor inhibition differentiates youths with Bipolar Disorder and Severe Mood Dysregulation. *Biol Psychol*. 2012;89(1):148-155.

42. De Los Reyes A, Kazdin AE. Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychol Bull*. 2005;131(4):483-509.

Table 1: Latent Class Analysis Model Fit Statistics

Number of Classes	Fit Statistics				
	-2 loglikelihood	**BIC**	**VLMR**	**BLRT**	**Entropy**
1	-39030.3	78443.79	NA	NA	1
2	-36850.3	74474.39	<0.001	<0.001	0.82
3	-36188.3	73541.19	<0.001	<0.001	0.80
4	-35664.2	72883.75	0.0046	<0.001	0.79
5	-35252	72450.08	0.0028	<0.001	0.81
6	-35067.2	72471.34	0.3346	<0.001	0.81
7	-34893.2	72514.15	0.7601	<0.001	0.81

Note: BIC = Bayesian Information Criterion; BLRT = Bootstrap Likelihood Ratio Test; NA = not applicable; VLMR = Vuong-Lo-Mendell-Rubin Likelihood Ratio.

Table 2: Strengths and Difficulties Questionnaire (SDQ) Summary Scores For Each Latent Class

Class (N)	SDQ Summary Scores				
	Emotional Symptoms (*M*, *SD*)	**Conduct Problems** (*M*, *SD*)	**Hyperactive Behavior** (*M*, *SD*)	**Peer Problems** (*M*, *SD*)	**Prosocial Behavior** (*M*, *SD*)
1 (373)	1.36, 1.30	1.98, 1.29	2.52, 1.51	1.66, 1.57	5.81, 1.36
2 (340)	4.20, 1.76	1.37, 1.19	3.16, 1.84	2.51, 1.71	8.99, 1.02
3 (261)	4.15, 2.23	4.32, 1.58	6.44, 1.98	2.91, 2.08	6.31, 1.97
4 (539)	1.20, 1.10	1.57, 1.17	3.99, 1.43	0.57, 0.80	8.59, 1.02
5 (613)	0.88, 1.05	0.64, 0.83	0.81, 0.87	0.89, 1.07	9.04, 1.03

Note: Class 3 comprised the "dysregulated" group who exhibited highest levels of impairment across all dimensions. Class 5 comprised the "low symptom" control group who exhibited the lowest levels of impairment. Summary scores were calculated by the sum of five items related to each dimension.[7] Higher values signify more difficulty except within the prosocial domain.

Table 3: Summary of Mediation Models

Model	*B*	*SE*	*p*	95% Bootstrapped CI	
				Lower	**Upper**
Anxiety to Dysregulation					
Total Effect	.690	.055	.001	.582	.799
Direct Effect	.667	.058	.001	.554	.780
Indirect Effect	**.023**	**.010**	**.025**	**.003**	**.043**
Depression to Dysregulation					
Total Effect	.700	.044	.001	.613	.788
Direct Effect	.683	.045	.001	.595	.772
Indirect Effect	.017	.009	.065	-.001	.035
Conduct Disorder to Dysregulation					
Total Effect	.780	.035	.001	.712	.848
Direct Effect	.762	.036	.001	.692	.833
Indirect Effect	**.018**	**.008**	**.022**	**.003**	**.033**
ODD to Dysregulation					
Total Effect	.841	.024	.001	.794	.888
Direct Effect	.826	.025	.001	.776	.876
Indirect Effect	.015	.008	.070	-.001	.031
ADHD to Dysregulation					
Total Effect	.919	.019	.001	.882	.957
Direct Effect	.909	.021	.001	.867	.951
Indirect Effect	.010	.006	.124	-.003	.023

Note: Total effects reflect association between the likelihood of disorder and dysregulation. Direct effects reflect the association between the likelihood of a disorder and dysregulation while accounting for the gray matter volume (GMV) region of interest (mediator). Indirect effects are the difference in betas, and reflect the magnitude of mediation through the region of interest. Significant indirect effects (95%CI >0) in bold. ADHD = attention-deficit/hyperactivity disorder; ODD = Oppositional Defiant Disorder; SE = standard error.

Figure Titles and Captions

Figure 1: Average Strengths and Difficulties Questionnaire (SDQ) Item Endorsement for Five Classes

Note: Each SDQ item present on the x-axis, ordered by the five respective SDQ domains to aid in interpretability. Average item endorsement on y-axis, from 0-2 (Not true, somewhat true, certainly true). Items with asterisks indicate reverse coding. Dysregulated class (3) in green line; low symptom class (5) in black line.

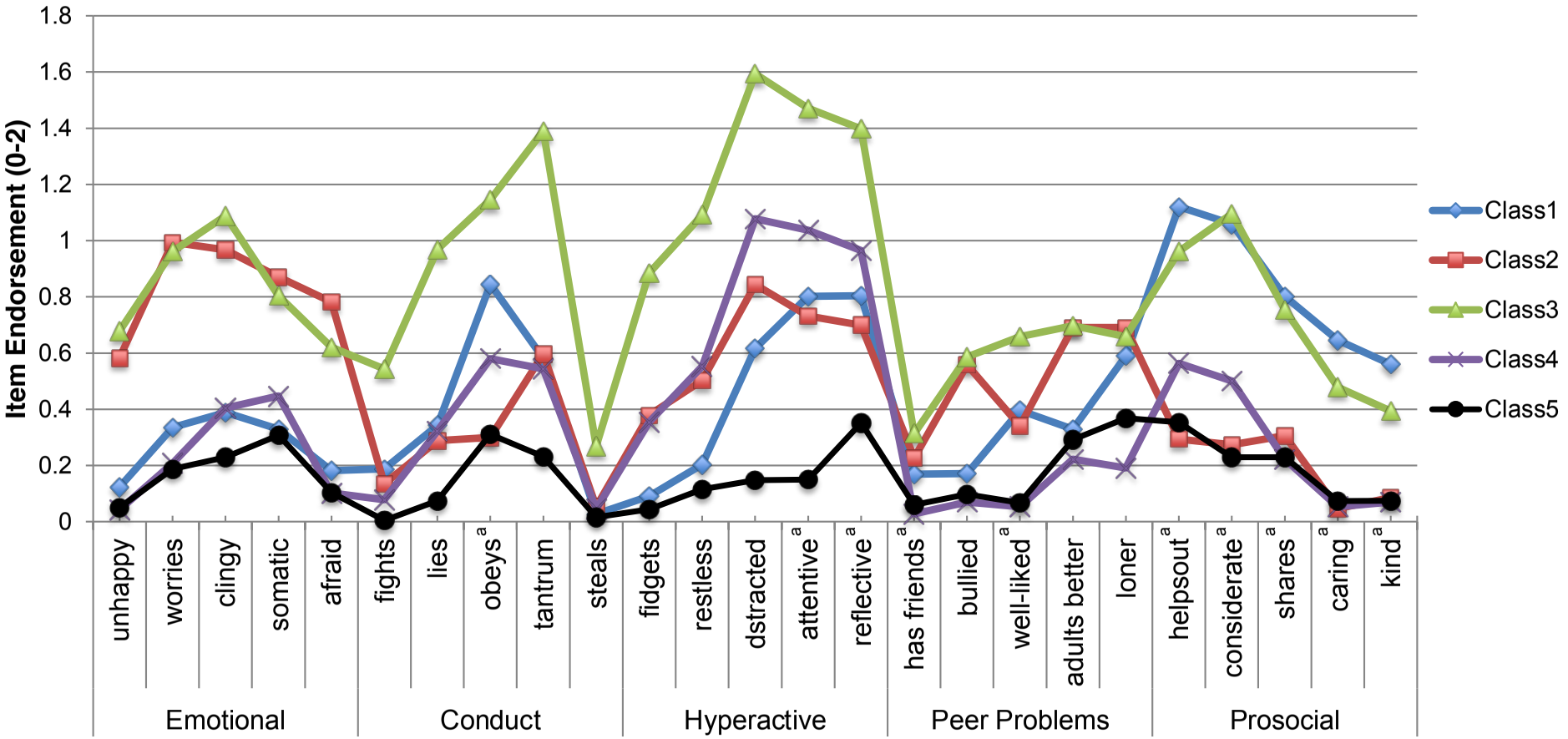
Figure 2: Right Orbitofrontal Cortex Region of Interest

Note: Cluster (k=48 voxels; center of mass Montreal Neurological Institute (MNI) coordinates: 3), identified as passing TFCE-correction ($p < .05$) from a two-sample residual gray matter volume permutation analysis. This cluster was also present in a two-group permutation analysis estimated without residualized images or nuisance covariates. TFCE = Threshold-Free Cluster Enhancement.[27]

Figure 3: Mediation Models with Significant Indirect Effects

Note: Path models of the relationship between the likelihood of anxiety disorder (left), or, conduct disorder (right), to dysregulation, mediated by the right orbitofrontal cortex gray matter volume region of interest (ROI). All coefficients are standardized and pass a null-hypothesis significance test at $p<.05$. The indirect effects (dotted line, c paths) reflect the magnitude of mediation through the ROI, with significance determined by 95% confidence intervals generated from 1000 bootstrapped samples (see Table 3). Total effects (c' paths) reflect the bivariate correlation between a disorder and dysregulation when the mediator is excluded. The negative parameter estimates for the paths into and out of the ROI (a and b paths) are in line with the lower gray matter volume (GMV) exhibited by the dysregulated group. OFC = orbitofrontal cortex.

Figure


Item Endorsement (0-2)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
unhappy
worries
clingy
somatic
afraid
fights
lies
obeys[a]
tantrum
steals
fidgets
restless
dstracted
attentive[a]
reflective[a]
has friends[a]
bullied
well-liked[a]
adults better
loner
helpsout[a]
considerate[a]
shares[a]
caring[a]
kind[a]
Emotional
Conduct
Hyperactive
Peer Problems
Prosocial
Class1
Class2
Class3
Class4
Class5

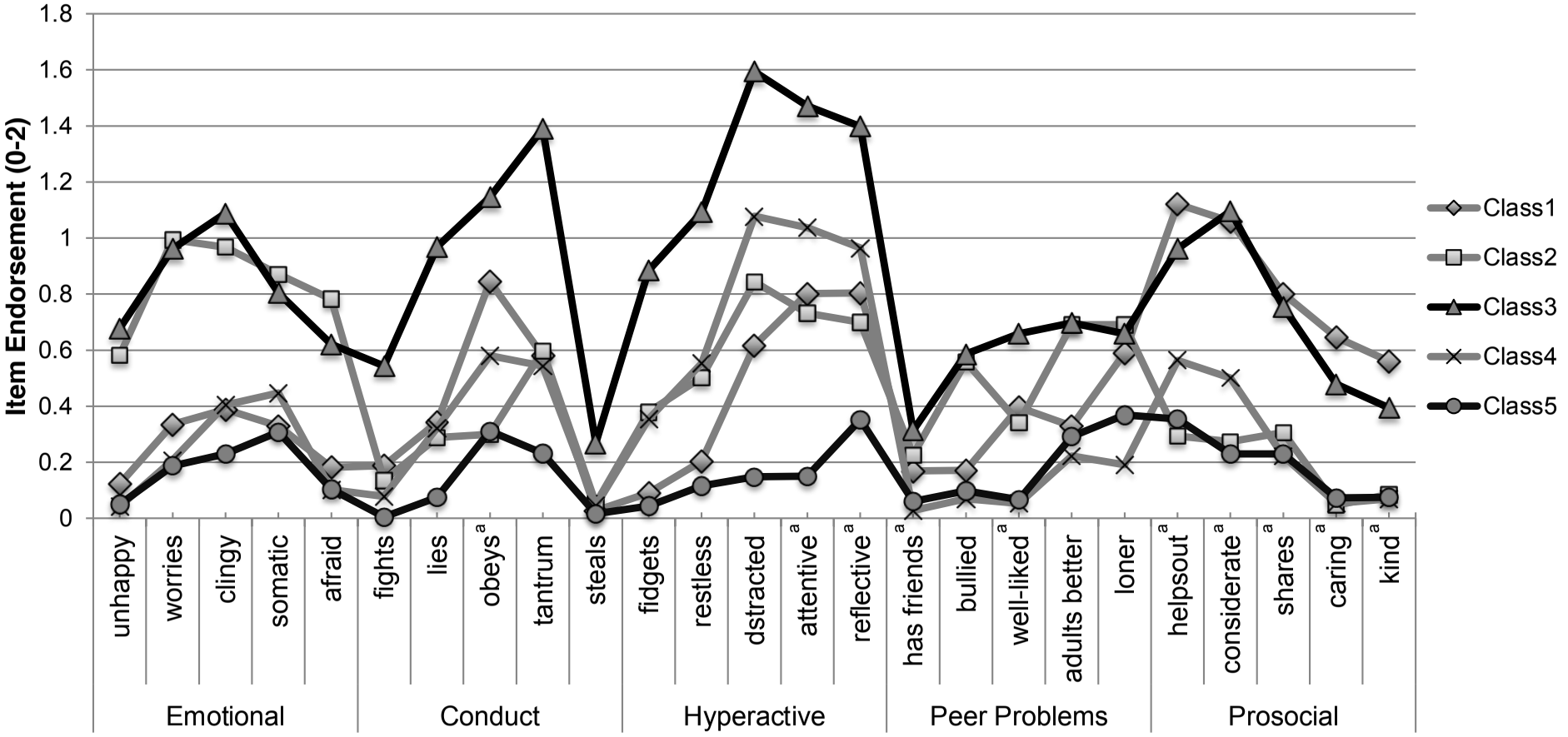
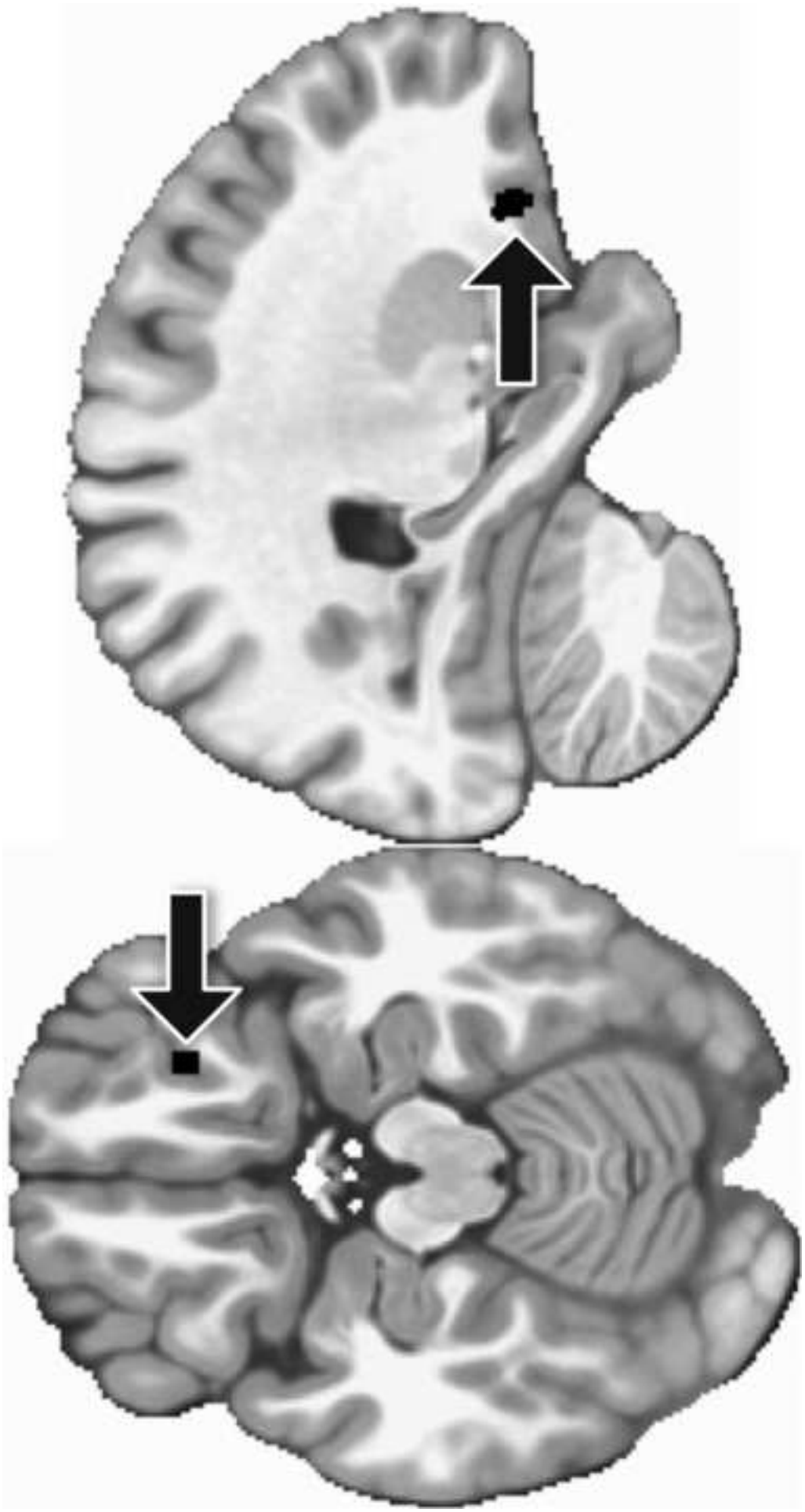
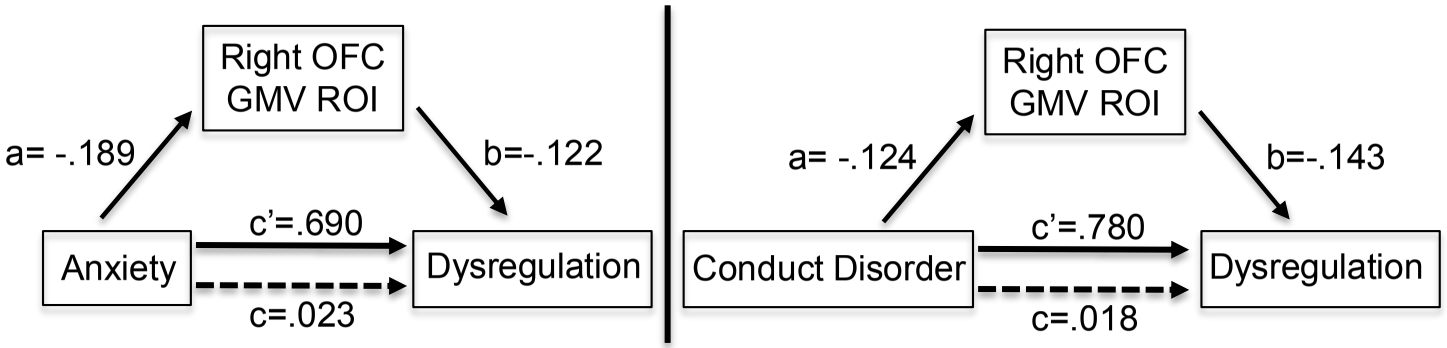

Item Endorsement (0-2)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
unhappy
worries
clingy
somatic
afraid
fights
lies
obeys[a]
tantrum
steals
fidgets
restless
dstracted
attentive[a]
reflective[a]
has friends[a]
bullied
well-liked[a]
adults better
loner
helpsout[a]
considerate[a]
shares[a]
caring[a]
kind[a]
Emotional
Conduct
Hyperactive
Peer Problems
Prosocial
Class1
Class2
Class3
Class4
Class5

Figure2



Figure

Figure3


Right OFC
GMV ROI
a= -.189
b=-.122
c'=.690
Anxiety
Dysregulation
c=.023
Right OFC
GMV ROI
a= -.124
b=-.143
c'=.780
Conduct Disorder
Dysregulation
c=.018

Supplemental 1

Materials and Methods

Structural MRI

High-resolution anatomical magnetic resonance images were acquired, including a 3D T1-weighted magnetization prepared gradient echo sequence based on the ADNI protocol (http://adni.loni.usc.edu). Structural MRI processing included data segmentation and normalization to the Montreal Neurological Institute template using the SPM optimized normalization routine. Gray matter images were modulated, thus facilitating comparisons of volumetric, rather than tissue concentration, differences.[1] N=2095 participants had data submitted to the morphometric processing pipeline.

Functional MRI

Full details of the magnetic resonance imaging (MRI) acquisition protocols and quality checks have been described previously, including an extensive period of standardization across MRI scanners.[2] MRI Acquisition Scanning was performed at the eight IMAGEN assessment sites (London, Nottingham, Dublin, Mannheim, Dresden, Berlin, Hamburg, and Paris) with 3T whole body MRI systems made by several manufacturers (Siemens: 4 sites, Philips: 2 sites, General Electric: 1 site, and Bruker: 1 site). To ensure a comparison of MRI data acquired on these different scanners, we implemented image acquisition techniques using a set of parameters compatible with all scanners that were held constant across sites, for example, those directly affecting image contrast or fMRI preprocessing.

For each task the following MRI acquisition parameters were used. 300 whole-brain volumes were collected using a gradient-echo T2*-weighted pulse sequence (EPI), with a TR=2,200ms, TE=30ms, matrix size=64x64; voxel resolution=3.4mm, with 40 slices collected in

descending order at a thickness of 2.4mm with 1mm spacing gap. The field of view was set parallel to the AC-PC line to minimize signal dropout.

Standardized hardware for visual and auditory stimulus presentation (NordicNeurolabs, Bergen Norway, http://www.nordicneurolab.com) was used at all sites. BOLD functional images were acquired with a gradient-echo echo planar imaging (EPI) sequence using a relatively short echo-time to optimize imaging of subcortical areas. Briefly, the functional imaging processing was as follows: Time series data were first corrected for slice-timing, then corrected for movement, non-linearly warped onto MNI space using a custom EPI template, and Gaussian-smoothed at 5mm-full width half maximum. Nuisance variables were also added to the design matrix: estimated movement was added in the form of 12 additional regressors (3 translations, 3 rotations, 3 translations shifted 1 TR before and 3 translations shifted 1 TR later). Each individual fMRI time series underwent automatic spike detection, using a mean-squared based metric to identify unexpected values temporally and spatially slice per slice. Time-points with artifacts (if any) of each sequence were regressed out of each participant's data by adding a corresponding number of regressors with value 1 at the time- point of the artifact and 0 elsewhere to the design matrix.

From the N=2024, with structural MRI data, there were n=1807 with stop task data, n=1814 with reward task data, and n=1889 with faces task data. Variability across fMRI sample sizes are due to many reasons, including but not limited to data quality control, technical errors during scanning, poor task performance, and incomplete scanning sessions.

Functional Tasks Information

Stop Signal Task (SST)

The SST required volunteers to respond to regularly presented visual go stimuli (arrows

pointing left or right) but to withhold their motor response when the go stimulus was followed unpredictably by a stop-signal (an arrow pointing upwards). Stopping difficulty was manipulated across trials by varying the delay between the onset of the go arrow and the stop arrow (stop-signal delay, SSD) using a previously described tracking algorithm.[3] A block contained 400 go trials and 80 variable delay stop trials with between 3 and 7 go trials between two stop trials. Stimulus duration in go trials was 1000 ms and in stop trials varied (0– 900ms in 50 ms steps) in accordance with the tracking algorithm (initial delay = 250 ms). We calculated contrast images for successful inhibitions ("stop success") and unsuccessful inhibitions ("stop fail"), both vs. an implicit baseline. Behavioral data from the stop signal task was incomplete due to technical errors, therefore this data was omitted from the modeling procedures, however, the stop signal task had an adaptive performance algorithm to account for individual differences in reaction time.

Monetary Incentive Delay (Reward Task)

The Monetary Incentive Delay (MID) task (adapted from a task described previously)[4] required participants to respond to a briefly presented target by pressing either a left-hand or right-hand button as quickly as possible to indicate whether the target appeared on the left or the right side of the monitor display. If the participants responded while the target was on the screen, they scored points but if they responded before the target appeared or after the offset of the target they received no points. A cue preceded the onset of each trial, reliably indicating the position of the target and the number of points awarded for a successful response. A triangle indicated no points (No Win), a circle with one line 2 points (Small Win) and a circle with three lines 10 points (Large Win). Twenty-two trials of each type were presented in a pseudo-random order. The duration of the target was adjusted adaptively so that 66% of the trials produced a correct

response. The participants were informed that at the end of the session they would receive one candy (M&M) for every five points won. We calculated contrast images for the anticipation period of Large Win minus No Win, and the outcome period for Large Win minus No Win.

Face Processing Task

The Face task involved passive viewing of video clips that displayed ambiguous (emotionally ‘‘neutral’’) or angry face expressions or control (nonbiological motion) stimuli.[5] Each trial consisted of short (2 to 5 s) black-and-white video clips depicting either a face in movement or the control stimulus. The control stimuli consisted of black-and-white concentric circles of various contrasts, expanding and contracting at various speeds, roughly matching the contrast and motion characteristics of the face clips. The stimuli were presented through goggles (Nordic Neurolabs, Bergen, Norway) in the scanner and subtended a visual angle of 10º by 7º. The video clips were arranged into 18-s blocks; each block included seven to eight video clips. Five blocks of each biological-motion condition (neutral and angry faces), and nine blocks of the control condition (circles) were intermixed and presented to the participant in a 6-minute run. We calculated contrast images from angry faces minus control stimuli, neutral faces minus control stimuli, and angry faces minus neutral faces. After the scanning session, participants completed a recognition task in which they were presented with three of the faces previously presented in the scanning session and two novel faces.

Head Motion

Head motion estimates were included as nuisance regressors twice— once during first level task activation analyses (12 regressors: 3 translation, 3 rotation, 3 translations shifted 1 TR before, and 3 translations shifted 1 TR later), and again during the between group ANCOVA models (mean framewise displacement). The framewise displacement (FD) for each participant

for each fMRI task was calculated using the six displacement parameters estimated during image realignment preprocessing procedures.[6] Based off prior developmental neuroimaging studies, a head motion exclusionary criterion of mean FD > .9mm was used.[7] For the stop signal task, 3 dysregulated participants were excluded. For the faces task, 1 dysregulated participant was excluded. For the MID task, 5 dysregulated and 1 low symptom participants were excluded. After applying this exclusion, mean FD for each subject for each task was then submitted to a 2-sample t-test to determine if head motion across each task significantly differed between the two groups. While mean FD was generally low for each group, there was still significantly greater head motion in the dysregulated group. See supplemental table 4 for results.

Additionally, we tested the hypothesis that head motion is associated with dysregulation status by running a correlation between the dysregulated class probability score (class 3 from the LCA) with the mean framewise displacement for each task. As expected, modest correlations were detected, suggesting that a higher probability of dysregulation class membership is associated with more head motion during the stop task (r=.149, p<.005), faces task (r=.141, p<.005), and MID task (r=.146, p<.005). As stated, this effect is expected as the two groups are known to exhibit between-group differences in head motion.

Characterizing Instruments

Wechsler Intelligence Scale for Children.

Participants completed a version of the Wechsler Intelligence Scale for Children WISC-IV,[8] of which we included the following subscales. Perceptual Reasoning, consisting of Block Design (arranging bi-colored blocks to duplicate a printed image) and Matrix Reasoning (in which a series of colored matrices are presented and the child is asked to select the consistent

pattern from a range of options). Verbal Comprehension consisting of Similarities (two similar but different objects or concepts are presented and the child is asked to explain how they are alike or different) and Vocabulary (a picture is presented or a word is spoken aloud by the experimenter and the child is asked to provide the name of the depicted object or to define the word).

Puberty Development Scale.

The Puberty Development Scale (PDS)[9] was used to assess the pubertal status of each participant. This scale provides an eight-item self-report measure of physical development based on the Tanner stages with separate forms for males and females. For this scale, there are five categories of pubertal status: 1= prepubertal, 2=beginning pubertal, 3=midpubertal, 4=advanced pubertal, 5=postpubertal. Participants answered questions about their growth in stature and pubic hair, as well as menarche in females and voice changes in males.

Missing Data

Dysregulated Group Missingness

From the starting N=233 dysregulated participants with an anatomical scan, 226 received a stop signal task scan. The reasons for missingness in those seven are unknown but are likely due to the participant electing to stop, scheduling conflicts, or scanner malfunctions. Of those 226, there were three participants excluded due to excessive head motion. Of the remaining 223, there were 37 participants who did not provide stop success images, and, 28 participants who did not provide stop failure images. Reasons for missing these contrast images may be due to any combination of reasons including poor task performance, technical errors, or poor image quality.

Hence, for the dysregulated group there were 186 participants with stop success images and 195 participants with stop failure images.

From the starting N=233 dysregulated participants with an anatomical scan, 222 received a reward task scan. The reasons for missingness in those 11 are unknown but are likely due to the participant electing to stop, scheduling conflicts, or scanner malfunctions. Of those 222, there were five participants excluded due to excessive head motion. Of the remaining 217, there were 22 who did not provide either reward anticipation or reward outcome images due to any combinations of reasons including poor task performance, technical errors, or poor image quality. Hence, for the dysregulated group there were 195 participants with both reward anticipation and reward outcome images.

From the starting N=233 dysregulated participants with an anatomical scan, 231 received a face processing scan. The reasons for missingness in those two are unknown but are likely due to the participant electing to stop, scheduling conflicts, or scanner malfunctions. Of those 231, there was one participant excluded due to excessive head motion. Of the remaining 230, there were 26 who did not provide any of the face processing contrast images due to any combinations of reasons including technical errors, or poor image quality. As this is a passive viewing task, behavioral performance quality control is not applicable. Hence, for the dysregulated group there were 204 participants with all face processing contrast images. See Table S3 for a breakdown of all fMRI contrast image sample sizes by group.

Low Symptom Group Missingness

From the starting N=233 low symptom participants with an anatomical scan, 228 received a stop signal task scan. The reasons for missingness in those five are unknown but are

likely due to the participant electing to stop, scheduling conflicts, or scanner malfunctions. No participants were excluded due to head motion. Of those 228, here were 34 participants who did not provide stop success images and 31 participants who did not provide stop failure images. Reasons for missing these contrast images may be due to any combination of reasons including poor task performance, technical errors, or poor image quality. Hence, for the low symptom group there were 194 participants with stop success images and 197 participants with stop failure images.

From the starting N=233 low symptom participants with an anatomical scan, 217 received a reward task scan. The reasons for missingness in those 16 are unknown but are likely due to the participant electing to stop, scheduling conflicts, or scanner malfunctions. Of those 217, there was one participant excluded due to excessive head motion. Of the remaining 216, there were 26 who did not provide either reward anticipation or reward outcome images due to any combinations of reasons including poor task performance, technical errors, or poor image quality. Hence, for the low symptom group there were 190 participants with both reward anticipation and reward outcome images.

From the starting N=233 low symptom participants with an anatomical scan, all 233 received a face processing scan and none were excluded due to head motion. Of those 233, there were 30 who did not provide any of the face processing contrast images due to any combinations of reasons including technical errors, or poor image quality. As this is a passive viewing task, behavioral performance quality control is not applicable. Hence, for the low symptom group there were 203 participants with all face processing contrast images. See Table S3 for a breakdown of all fMRI contrast image sample sizes by group.

Additional Analyses

Balancing of Site

While the dysregulated and low symptom groups were best matched on site, we tested whether any non-significant differences in site may have influenced the findings. To do so, we randomly sampled without replacement to perfectly balance the two groups on each site. From the starting n=466, a random sample of n=408 was identified (n=204 for each group), such that each group had a perfectly balanced representation from each of the eight sites. A two-sample t-test yielded consistent findings with the original sample, such that lower GMV was found in the dysregulated group (t_{406} = -3.34, p < .001). Therefore, we do not suspect the non-significant difference in site distribution to be driving the overall finding of the paper.

Random Subsampling Test to Balance Head Motion

While head motion estimates were accounted for in the first-level contrast image estimation, and, again as a nuisance covariate (mean FD) in the between group comparison, we note that the dysregulated group had significantly higher head motion than the low symptom group, and partialling out these differences does not negate the significant difference. To assess whether head motion was driving the between-group differences for stop success, a random subsampling procedure was implemented. Here, we identified two random subsamples of the dysregulated and low symptom group who failed to differ on head motion. To do this, we randomly sampled without replacement 80% of the dysregulated sample, and 83% of the low symptom. These percentages were chosen to use consistent subsample rates of the fMRI data relative to the anatomical data (See Table S3).

A random subsample of n=149 dysregulated and n=161 low symptom participants were selected until a two-sample *t*-test failed to identify between-group head motion differences (p >

.05). Next, these pseudo-random subsamples were submitted to a similar ANCOVA model, regressing the stop success activations against the two groups, accounting for head motion (mean FD). This procedure was performed 5,000 times, while saving the *F*- and *p*-value of the dysregulation status coefficient (the between-subject effect). This procedure resulted in a mean $F_{1,307}$=4.9, *p*=.045. As consistent effects were found on these pseudo-random samples, these results indicate that the between-group differences in stop success activation persist after effectively removing the differences in head motion.

Supplement References

1. Ashburner J, Friston KJ. Voxel-Based Morphometry—The Methods. *NeuroImage*. 2000;11(6):805-821.

2. Schumann G, Loth E, Banaschewski T, et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol Psychiatry*. 2010;15(12):1128-1139.

3. Rubia K, Smith AB, Brammer MJ, Toone B, Taylor E. Abnormal Brain Activation During Inhibition and Error Detection in Medication-Naive Adolescents With ADHD. *Am J Psychiatry*. 2005;162(6):1067-1075.

4. Knutson B, Fong GW, Adams CM, Varner JL, Hommer D. Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*. 2001;12(17):3683-3687.

5. Grosbras M-H. Brain Networks Involved in Viewing Angry Hands or Faces. *Cereb Cortex*. 2005;16(8):1087-1096.

6. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 2012;59(3):2142-2154.

7. Silvers JA, Lumian DS, Gabard-Durnam L, et al. Previous Institutionalization Is Followed by Broader Amygdala–Hippocampal–PFC Network Connectivity during Aversive Learning in Human Development. *J Neurosci*. 2016;36(24):6420-6430.

8. Wechsler D. Wechsler intelligence scale for children–Fourth Edition (WISC-IV). *San Antonio TX Psychol Corp*. 2003.

9. Carskadon MA, Acebo C. A self-administered rating scale for pubertal development. *J Adolesc Health*. 1993;14(3):190-195.

Table S1: Dysregulated Class Participants Compared to All Other Latent Class Analysis Participants

Measure	Group		*p*
	Dysregulation (N=258)	**All Other Groups (N=1856)**	
Sex (Male, Female)	146, 112	887, 969	.01
Age in years (*M, SD*)	14.56, .43	14.55, .44	.50
Performance IQ (*M, SD*)	103.59, 13.97	108.01, 13.64	.01
Verbal IQ (*M, SD*)	106.83, 14.38	111.37, 13.56	.01
Pubertal Development (*M, SD*)	2.83, .55	2.92, .56	.01
Handedness (Right, Left)	235, 23	1669, 187	.56
London (N)	29	240	.42
Nottingham (N)	41	246	.27
Dublin (N)	32	203	.51
Berlin (N)	38	230	.31
Hamburg (N)	32	234	.90
Mannheim (N)	25	239	.14
Paris (N)	31	232	.80
Dresden (N)	30	232	.66

Note: Comparison of participant characteristics for those identified from the latent class analysis. The dysregulation class (Class 3) compared to the individuals from all other classes. Relative to main text Table 2, there were three dysregulated participants and nine other class participants with incomplete demographic information.

Table S2: Dysregulated and Low Symptoms Group Characteristics

Measure	Groups		*p*
	Dysregulation (N=233)	**Low Symptoms (N=233)**	
Sex (Male, Female)	127,106	127,106	1.00
Age in years (*M, SD*)	14.58, .45	14.55, .45	.60
Performance IQ (*M, SD*)	103.48, 14.56	103.20, 12.70	.82
Verbal IQ (*M, SD*)	106.34, 14.88	106.85, 11.70	.69
Pubertal Development (*M, SD*)	2.82, .54	2.86, .57	.42
Total Gray Matter Volume mm^3 (*M, SD*)	729.92, 69.56	737.56, 65.35	.18
Handedness (Right, Left)	(212, 22)	(211, 21)	.87

Note: Dysregulated individuals identified from a latent class of respondents most likely to endorse problems related to prosocial behaviors, attention, positive affect, and conduct. The comparison group ("low symptoms") were identified from a latent class of respondents who were least likely to endorse similar problems, and randomly sampled until a subsample of individuals failed to exhibit between-group differences on sex, age, performance and verbal IQ, pubertal development, total gray matter volume, and handedness. Total Gray Matter Volume was calculated from the whole-brain VBM image. All other tabulated features were calculated from self-report measures, with IQs measured via the Weschler Intelligence scale and Pubertal Development measured via the PDS (described above). The PDS measure is an average score (from 0-4) across five items detailing physiological changes in males and females separately.

Table S3: Outline of Samples Used for Each fMRI Comparison

fMRI Contrast	Dysregulated Class		Low Symptom Class	
	Analytic Sample N, %	**Lost Sample N, %**	**Analytic Sample N, %**	**Lost Sample N, %**
Stop Success	186 (.80)	37 (.16)	194 (.83)	34 (.15)
Stop Fail	195 (.84)	28 (.12)	197 (.85)	31 (.13)
Neutral Faces	204 (.88)	26 (.11)	203 (.87)	30 (.13)
Angry Faces	204 (.88)	26 (.11)	203 (.87)	30 (.13)
Angry-Neutral Faces	204 (.88)	26 (.11)	203 (.87)	30 (.13)
Reward Anticipation	195 (.84)	22 (.09)	190 (.82)	26 (.11)
Reward Outcome	195 (.84)	22 (.09)	190 (.82)	26 (.11)

Note: Only participants who received the fMRI scan were considered as candidates for lost data, hence, the analytic and lost sample sizes may not sum to the 233 that received an anatomical scan.

Table S4: Analysis of Head Motion in Reduced Samples with fMRI data

Task	Framewise Displacement						Dysregulated vs. Low Symptom
	Dysregulation Group			Low Symptom Group			
	N	*M*	*SD*	*N*	*M*	*SD*	
Stop Signal	186	.14	.10	194	.11	.08	$t_{378}= 2.9$, $p < .05$
Faces	204	.13	.12	203	.10	.08	$t_{405}= 2.7$, $p < .05$
MID	195	.18	.14	190	.14	.11	$t_{383}= 3.1$, $p < .05$

Note: MID = Monetary Incentive Delay Task.